Random Projection Information Splits for Reduced-Sensitivity Nonlinear Processing

Abstract

A useful way to view random projection systems is not primarily as dimensionality reduction methods, but as *linear algebra information splitting systems*. The input signal is transformed into multiple partially overlapping views, where each view contains only a constrained portion of the total informational complexity of the original data.

Individually, these views have lower sensitivity to perturbation and noise. The real benefit appears when nonlinear functions or neuron-like processing elements operate on the split representations rather than on the original full input.

Each nonlinear unit therefore processes:

- a reduced-complexity view,
- a lower-sensitivity manifold slice,
- and a partially filtered version of the original information space.

This may substantially reduce nonlinear instability and nuisance sensitivity while preserving useful large-scale structure.

---

1. Information Splitting via Random Projection

Consider an input vector:

$$x \in \mathbb{R}^N$$

Apply a full-dimensional random mixing transform, for example a Walsh-Hadamard random projection:

$u = HDx$

where:

- H is the Walsh-Hadamard transform,
- D is a random sign matrix.

The dimensionality itself is unchanged. The transform simply redistributes information globally across the representation.

Next, generate multiple subsampled views:

$u_i = S_i u$

where each S_i selects a subset of transformed coefficients.

Each view is then reconstructed back into the original domain:

$y_i = DHS_i^T S_i HDx$

Each reconstructed y_i:

- preserves some information,
- suppresses other information,
- and represents a partial low-sensitivity view of the original signal.

---

2. Reduced Sensitivity in Each Split

Each reconstruction only contains a subset of the total transformed information.

As a result:

- many perturbation directions vanish,
- high-frequency instability is reduced,
- nuisance variation weakens,
- and local sensitivity decreases.

Importantly, this is not simple coordinate dropout. The information remains globally distributed through the transform mixing process.

Each split therefore retains:

- coherent large-scale structure,
- distributed semantic information,
- and manifold-supported variation

while reducing sensitivity to small local perturbations.

---

3. Why Nonlinear Processing Benefits

A standard neural layer operates directly on the full input:

$$h = \sigma(Wx)$$

Every nonlinear unit must therefore handle:

- the full perturbation complexity,
- all nuisance dimensions,
- all high-frequency variation,
- and all unstable feature interactions simultaneously.

This can produce:

- chaotic activation geometry,
- brittle decision boundaries,
- and excessive sensitivity to noise.

Split-View Processing

Now instead suppose nonlinear units operate on multiple reconstructed views:

$$h_i = (W_i y_i)$$

Each nonlinear unit now sees only:

- a partial information view,
- a reduced-complexity representation,
- and a lower-sensitivity signal geometry.

The nonlinear function therefore operates on:

- smoother trajectories,
- more stable perturbation structure,
- and less informational “noise.”

The key idea is:

> The linear projection split reduces the informational burden seen by each nonlinear element.

---

4. The Role of Subsample Size

The subsample size controls the balance between:

- fidelity,
- sensitivity,
- smoothness,
- and robustness.

Small Subsamples

Very strong filtering:

- low sensitivity,
- high smoothing,
- reduced instability,
- but less preserved detail.

Large Subsamples

Higher fidelity:

- more preserved information,
- greater expressiveness,
- but increased sensitivity.

Intermediate Regime

This regime is especially interesting because it may preserve:

- useful manifold structure,
- semantic organization,
- and dominant correlations

while still suppressing:

- nuisance perturbations,
- unstable fluctuations,
- and excessive nonlinear sensitivity.

---

5. Multiple Covering Views

Using many overlapping projection views creates:

- multiple partial manifold observers,
- distributed low-sensitivity representations,
- and overlapping geometric decompositions of the input.

Collectively, the views preserve all important information while allowing each nonlinear processing stage to operate on a simpler and more stable signal geometry.

---

6. Conclusion

Random projection information splitting provides an interesting alternative perspective on representation learning.

The important operation is not merely dimensional reduction, but:

- decomposition of information into multiple lower-sensitivity views.

Each reconstructed split:

- preserves partial global structure,
- suppresses some perturbation directions,
- and reduces the informational complexity presented to downstream nonlinear functions.

The primary benefit appears after nonlinear processing:

- each neuron or nonlinear unit no longer has to operate on the full informational noise of the original input space.

Instead, it processes:

- smoother,
- partially constrained,
- lower-sensitivity manifold views

which may improve:

- robustness,
- optimization stability,
- nonlinear smoothness,
- and distributed representation quality.